A COMPARISON OF TOURNAMENTS

by

© MICHAEL J.D. HOPKINS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

THE UNIVERSITY OF ALBERTA

EDMONTON, ALBERTA

SPRING, 1969

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "A COMPARISON OF TOURNAMENTS", submitted by MICHAEL J.D. HOPKINS in partial fulfillment of the requirements for the degree of Master of Science.

ABSTRACT

A class of random knock-out tournaments has been introduced by Narayana and Zidek [6], which is of interest in providing several feasible alternatives to round-robin tournaments under very general conditions. In this thesis we study some of the combinatorial properties of these tournaments and with the help of a computer obtain several results useful in the comparison of random knock-out tournaments. A brief review of procedures for the selection of the best object with (at least) a predetermined probability P' is also given.

ACKNOWLEDGEMENTS

I am indebted to Dr. T.V. Narayana for suggesting to me the topic and for his patient and enthusiastic guidance during the research and preparation of this thesis. I should also like to thank Dr. K.W. Smillie for his useful suggestions in preparing some of the programs in the Appendix, and to Miss June Talpash for typing the thesis.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The combinatorics of round-robin tournaments has often been studied in the literature (cf. David [2], Moon [5]), but relatively little work has been published on random knock-out tournaments. Recently Narayana and Zidek [6] have introduced a very large and interesting class of random knock-out tournaments and have studied their combinatorial properties when a stronger player is present. Their work on statistical inference in random tournaments suggests, in certain special cases at least, that random knock-out tournaments provide a convenient alternative to round-robin tournaments in selecting the best object amongst a group of N objects. It is our purpose to study further the properties of this class of random knock-out tournaments, and to provide a comparison of these tournaments when one, then two, stronger players are present (all other players are assumed to be of equal strength). Such a comparison of tournaments provides valuable insight in providing efficient procedures for the selection of the best object with a predetermined probability P'.

In the second chapter we give a summary of preliminary results useful in comparing tournaments, then introduce some results on the number of random knock-out tournaments of a

certain length, and a rule for a rough ordering of random knock-out tournaments.

It is complicated and perhaps undesirable to compare every single tournament against another, and as the thesis progresses it will be clear that no attempt is made to do this for the general case when there are $N \geq 4$ players present. A natural basis in comparing random knock-out tournaments is to say that a particular "tournament pattern" is preferred to another if and only if the probability of the strong player winning (i.e. playing in and defeating his opponent in the final round) is greater in this "pattern" than in the other (for a fixed number of players). The major theorem of Chapter III provides a comparison of several important "tournament patterns" which have been discussed in the literature (both recent and past) on random knock-out tournaments, including the classical case when $N = 2^t$ $(t = 2,3,...)$ players. At the end of Chapter III we extend the model with one strong player present to the model where two strong players of equal strength are present, which, it is hoped, will serve as a check when selecting the best object out of a group of N objects in simulation problems.

A systematic comparison of "tournament patterns" would have been handicapped due to the large amount of tedious calculation involved in providing explicit numerical results, if it were not for the use of a high speed digital computer.

In particular the programming language APL (A Programming Language-developed by Iverson [4]) provided a convenient medium in this context, because of the ease in learning the language and the precision with which unwieldy expressions could be formalized and then evaluated.

A brief review of the work by David [2] in the selection of a subset containing the best object is given in the fourth chapter. Applying his results we then give an idea how repeated knock-out tournaments compare to round-robin tournaments when one, then two strong players are present, by simulation using a table of random digits. Finally we provide an Appendix listing APL programs and tables which were useful in the comparisons.

CHAPTER II

PROCEDURES USEFUL IN COMPARING TOURNAMENTS

§2.1 Summary of Results

The following results which form the basis for a comparison of tournaments were first announced by Narayana and Zidek [6]. They defined a class of random knock-out tournaments (or tournaments) as follows:

Definition 2.1.1

For every integer $N \geq 2$, a class of random knock-out tournaments with N players is defined as a vector of positive integers $(m_1, \ldots, m_k)$ satisfying the following conditions:

(a) $m_i \geq m_k = 1$ is an integer for $i = 1, \ldots, k-1$.

(b) $m_1 + \ldots + m_k = N-1$.

(c) $2m_1 \leq N$,

$2m_i \leq N - m_1 - \ldots - m_{i-1} \quad (i \geq 2)$.

A tournament defined by the vector $(m_1, \ldots, m_k)$ is played as follows. In the first round $2m_1$ players, chosen at random from N, are paired off randomly. The remaining $N-2m_1$ players have a "bye" for this round. The m_1 pairs yield m_1 losers who are eliminated from the tournament. We are then

left with a tournament of $(N-m_1)$ players, with the vector $(m_2,\ldots,m_k)$. This inductive rule is well defined for $N > 2$, since in the case $N = 2$, there is a unique tournament of one round.

In order to prove the main theorem of Chapter III, we shall consider the following vectors, then reproduce a theorem first proved by Narayana and Zidek [6], concerning the probability that a strong player wins the tournament with vector $(m_1,\ldots,m_k)$.

Let

$$\underline{N} = (N_1,N_2,\ldots,N_k) \tag{2.1.1}$$

where $N_1 = N$, $N_i = N_{i-1} - m_{i-1}$ $(i \geq 2)$, and let

$$\underline{p} = (p_1,p_2,\ldots,p_k) \tag{2.1.2}$$

where $p_i = \dfrac{2m_i}{N_i}$, $q_i = 1 - p_i$ $(i = 1,\ldots,k)$.

Clearly N_i is the number of players in the i^{th} round, so that p_i is the probability that a specified player among these N_i does not obtain a bye in round i. We also note that $N_k = 2$ from Definition 2.1.1 (a), and hence $p_k = 1$.

We emphasize that we are studying the model where one player A has probability p of defeating each of the $(N-1)$ players $B_1,\ldots,B_{N-1}$, who are assumed to be of equal strength. It is clear that there exists a relation between the probability Π, that A wins the tournament, and the

probability Π^* that one of the $(N-1)$ B_j's wins the tournament, namely

$$\Pi + (N-1)\Pi^* = 1 \quad . \qquad (2.1.3)$$

Thus once Π is known, Π^* is immediately determined, leading us to the following theorem:

Theorem 2.1.1

The probability $\Pi = \Pi_N$, that A wins the tournament with vector $(m_1,\ldots,m_k)$, is given by :

$$\Pi = (p_1p+q_1)(p_2p+q_2)\ldots(p_{k-1}p+q_{k-1})p \quad . \qquad (2.1.4)$$

Proof

The theorem will be proved by induction on N . It is easily shown that the theorem is true for all tournaments when $N = 2,3,4$ players, by direct calculation. Assuming it is true for a number of players less than or equal to $N-1$, we now proceed inductively and prove it for N players.

A tournament of N players is defined by its vector $\underline{m} = (m_1,m_2,\ldots,m_k)$. The player A can survive the first round in the following two exclusive and exhaustive ways:

(a) he plays the first round and wins with probability $p \times \frac{2m_1}{N} = pp_1$,

(b) he has a bye with probability $\frac{N-2m_1}{N} = q_1$.

From the theorem of total probabilities, the player A survives the first round with probability,

$$p_1 p + q_1 \quad . \tag{2.1.5}$$

We are now in the reduced case of a tournament of $N-m_1 = N_2$ players with vector $(m_2,\dots,m_k)$. Since $N_2 \leq N-1$ we can apply the inductive hypothesis that:

$$\Pi_{N_2} = \Pi_{N_2;m_2,\dots,m_k} = (p_2 p+q_2)\dots(p_{k-1}p+q_{k-1})p \, , \tag{2.1.6}$$

where the notation of the L.H.S. is obvious.

The probability Π_N, that A wins the tournament, is given using (2.1.5) and (2.1.6) by:

$$\Pi_N = (p_1 p+q_1)\Pi_{N_2} \, , \tag{2.1.7}$$

thus completing the proof of Theorem 2.1.1.

On consideration of the expression (2.1.4), it is clear that the coefficients of the p^i $(i = 1,2,\dots,k)$ are the probabilities that a particular player (not necessarily the strong player A), receives k-1, k-2, ..., 0 byes in winning the tournament. If these probabilities are denoted by the vector $\underline{b}' = (b_{k-1}, b_{k-2},\dots,b_o)$, and if the vector $\underline{w}' = (p,p^2,\dots,p^k)$, then Π may be written as an inner product of two vectors:

$$\Pi = \underline{b}'\underline{w} \quad . \tag{2.1.8}$$

Once Π is known for any tournament with vector $(m_1,...,m_k)$, the expected number of games $E(R)$ played by A in the tournament may be calculated using the following theorem:

Theorem 2.1.2

$$\Pi + q\,E(R) = 1 \quad , \tag{2.1.9}$$

where $q = 1-p$.

The proof is similar to that of Theorem 2.1.1 using induction on N . This theorem is of help when such considerations as cost per game are included in the comparison of tournaments.

The program *VECTOR* generates all tournament vectors for a fixed $N > 4$, the programs *PROB* and *BYE* evaluate the expressions (2.1.4) and $\underline{b}'$ (the bye vector) of (2.1.8) respectively. These programs are available in the Appendix, (Programs 1, 2 and 4 respectively).

§2.2 The Number of Tournaments of Length k

It is perhaps undesirable to provide a rule for comparing every single tournament satisfying Definition 2.1.1 when there are N players present, due to the large number of different tournaments involved. Capell and Narayana [1] have obtained bounds for the number of tournaments T_N when there are N players present, viz.

$$\frac{160}{256} \cdot 2^{N-3} < T_N \leq \frac{165}{256} \cdot 2^{N-3} , \qquad N \geq 11 , \tag{2.2.1}$$

and have presented the following table of values for T_N when $N \leq 11$,

Table 2.2.1

N	2	3	4	5	6	7	8	9	10	11
T_N	1	1	2	3	6	11	22	42	84	165

In this section we shall provide an upper bound for the number of tournaments of length k . If $T(N,k)$ represents the number of tournament vectors $(m_1,\ldots,m_k)$ of length k, where $N = 2^t + K$ players $(t = 1,2,\ldots; 0 \leq K < 2^t)$, then clearly,

$$T(N,k) = 0 \quad \text{for} \quad k \leq t ,$$
$$\text{except for} \quad T(N,t) = 1 \quad \text{when} \quad N = 2^t . \tag{2.2.2}$$

Particular values of $T(N,k)$ may be obtained, by noting that

$$T(N,k) = T(N-1,k-1) + T(N-2,k-1) + \ldots + T([\tfrac{N+1}{2}],k-1) .$$

When N is even $(= 2n)$,

$$T(N,k) = T(n,k-1) + T(n+1,k-1) + \ldots + T(2n-1,k-1) .$$

If N is odd $(= 2n-1)$,

$$T(N,k) = T(n,k-1) + T(n+1,k-1) + \ldots + T(2n-2,k-1),$$

$$\therefore\ T(2n,k) = T(2n-1,k)+T(2n-1,k-1) \quad . \qquad (2.2.3)$$

Similarly,

$$T(2n+1,k) = T(2n,k)+T(2n,k-1)-T(n,k-1) \quad . \qquad (2.2.4)$$

Using (2.2.3) and (2.2.4) we obtain a table of values for $T(N,k)$.

Table 2.2.2

N \ k	2	3	4	5	6	7	8	9	10
3	1								
4	1	1							
5		2	1						
6		2	3	1					
7		1	5	4	1				
8		1	6	9	5	1			
9			6	15	14	6	1		
10			6	21	29	20	7	1	
11			4	26	50	49	27	8	1

By comparing this table with Pascal's Triangle of Binomial Coefficients, we obtain an upper bound for $T(N,k)$:

Lemma 2.2.1

$$T(N,k) \leq \binom{N-3}{k-2} , \qquad (k \geq 2;\ N \geq 3) \quad . \qquad (2.2.5)$$

Proof: (by induction on N)

It is clearly true for $N = 3,4,5,6$ on consideration of Table 2.2.2.

We assume the hypothesis is true for N odd $(= 2n-1)$, $n = 2,3,\ldots$.

Hence,

$$T(2n-1,k) \leq \binom{2n-4}{k-2} , \quad \text{and} \quad T(2n-1,k-1) \leq \binom{2n-4}{k-3} .$$

And,

$$T(2n,k) = T(2n-1,k)+T(2n-1,k-1) , \quad \text{by (2.2.3)}$$

$$\leq \binom{2n-4}{k-2} + \binom{2n-4}{k-3} ,$$

$$\therefore \quad T(2n,k) \leq \binom{2n-3}{k-2} , \tag{2.2.6}$$

(cf. Feller [3], p. 49) ,

and our hypothesis is true for N even .

$$T(2n+1,k) \leq T(2n,k) + T(2n,k-1) ,$$

from (2.2.4), since $T(n,k-1) \geq 0$.

Hence, $$T(2n+1,k) \leq \binom{2n-3}{k-2} + \binom{2n-3}{k-3} , \quad \text{by (2.2.6)} .$$

$$\therefore \quad T(2n+1,k) \leq \binom{2n-2}{k-2} , \tag{2.2.7}$$

and our hypothesis is true for N odd.

Using (2.2.6) and (2.2.7) we have,

$$T(N,k) \leq \binom{N-3}{k-2} , \quad (k \geq 2;\ N \geq 3) , \tag{2.2.8}$$

which completes the proof of the lemma.

We note that,

$$T_N = \sum_{k=2}^{N-1} T(N,k) \; , \qquad (N \geq 3)$$

$$\leq \binom{N-3}{0} + \binom{N-3}{1} + \; \ldots \; + \binom{N-3}{N-3} \quad \text{, by Lemma 2.2.1 .}$$

Using the result in Feller ([3], p. 50), we have

$$T_N \leq 2^{N-3} \; , \qquad \text{for} \quad N \geq 3 \; , \qquad (2.2.9)$$

providing an upper bound for T_N (although not as good as in (2.2.1)).

§2.3 Descendant Principle

In this section we introduce the idea of a descendant of a tournament, and show how it may be used in providing a reasonable ordering of tournament vectors ,with respect to the comparison described in Chapter I.

Definition 2.3.1

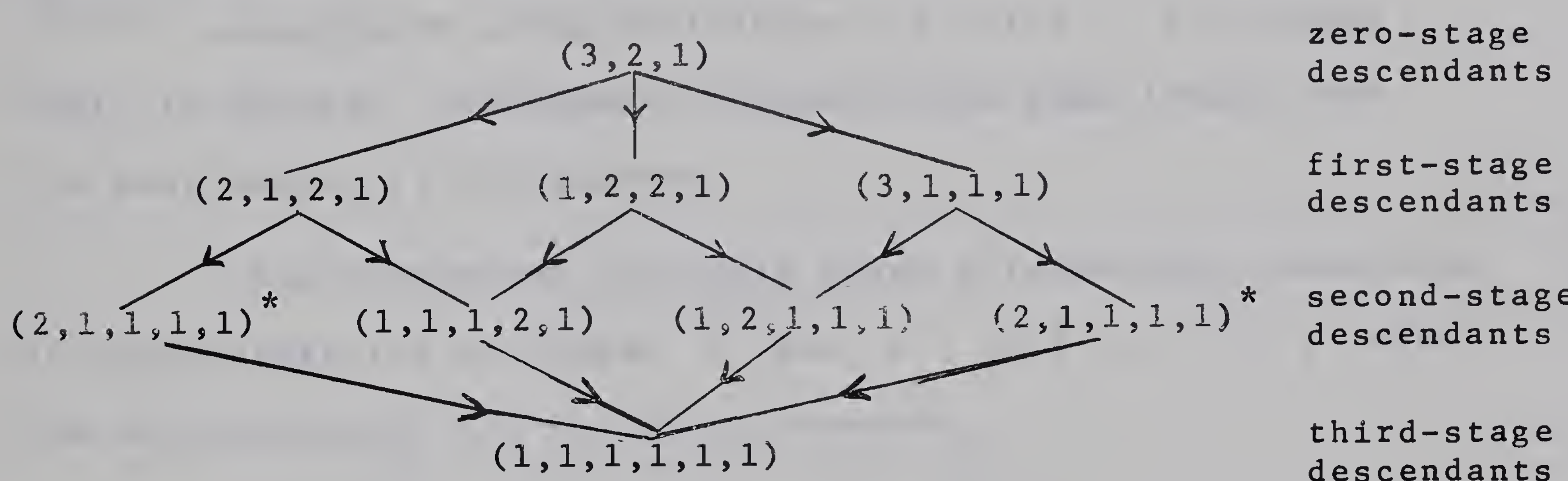
If $\underline{T} = (m_1, m_2, \ldots, m_k)$ is a tournament vector for N players, and for some $i < k$, $m_i = m_i' + m_i''$, where m_i', m_i'' are positive integers; then the tournament vector $\underline{T}' = (m_1, \ldots, m_{i-1}, m_i', m_i'', m_{i+1}, \ldots, m_k)$ will be called a <u>first-stage descendant</u> of $\underline{T}$.

<u>Note</u>: $\underline{T}$ may be considered a zero-stage descendant of itself; also if $\underline{T}''$ is a first-stage descendant of $\underline{T}'$, it is called

a second-stage descendant of $\underline{T}$. We continue in this way, defining the i^{th}- stage $(i = 0,1,2,\ldots)$ descendant of $\underline{T}$, until we arrive at the vector $(1,1,\ldots,1)$, which has no descendants (except itself). We shall use the term "descendants of $\underline{T}$" to denote all tournament vectors which are the i^{th}-stage descendants of $\underline{T}$ (including $\underline{T}$ itself).

Example

The descendants of $\underline{T} = (3,2,1)$.



* - the same vector.

Remarks

(1) It is apparent that $(1,1,1,1,1,1)$ is a descendant of every vector.

(2) From Table 2.2.1 we see that there are 11 tournament vectors of the type $(m_1,\ldots,m_k)$ when $N = 7$, thus the vector with the least number of elements does not necessarily have all the possible vectors as descendants. In this case there are three missing, namely

(2,2,1,1); (1,3,1,1); (1,1,2,1,1) .

(3) We notice further that the vectors of the same length have the same number of descendants. In fact the number of descendants of a vector, including itself, is in general 2^{N-k-1}. This is because any positive integer m_j, say, has 2^{m_j-1} compositions, (cf. Feller [3], p. 37). Clearly, each integer of the vector $\underline{m} = (m_1, \ldots, m_k)$ has $2^{m_1-1}, 2^{m_2-1}, \ldots, 2^{m_k-1}$ compositions respectively. Hence, the vector $\underline{m}$ has $2^{m_1-1} \cdot 2^{m_2-1} \cdot \ldots \cdot 2^{m_k-1} = 2^{(m_1+m_2+\ldots+m_k)-k} = 2^{N-k-1}$ descendants using Definition 2.1.1 (b). It follows that, in general, tournament vectors of the same length have the same number of descendants.

The descendant principle gives a reasonable comparison of tournaments for any fixed N and p, $(0.5 < p < 1)$, with the application of the following theorem:

<u>Theorem 2.3.1</u>

If $\underline{T}'$ is a first-stage descendant of $\underline{T}$, then

$$\Pi_{\underline{T}} > \Pi_{\underline{T}'} , \tag{2.3.1}$$

for a fixed N and p, $(0.5 < p < 1;\ N \geq 4)$, where $\Pi_{\underline{T}}$ = probability that player A wins tournament $\underline{T}$.

<u>Note</u>: If $\underline{T} = (1,1,\ldots,1)$, it has no descendants, and the theorem cannot be applied.

<u>Proof</u>

Using Theorem 2.1.1,

$$\Pi_{\underline{T}} = \Pi_{(m_1,\ldots,m_k)} ,$$
$$= (p_1p+q_1)\ldots(p_{i-1}p+q_{i-1})(p_ip+q_i)(p_{i+1}p+q_{i+1})\ldots(p_{k-1}p+q_{k-1})p ,$$
$$(2.3.2)$$

and using Definition 2.3.1,

$$\Pi_{\underline{T}'} = \Pi_{(m_1,\ldots,m_{i-1},m_i',m_i'',m_{i+1},\ldots,m_k)} ,$$
$$= (p_1p+q_1)\ldots(p_{i-1}p+q_{i-1})(p_i'p+q_i')(p_i''p+q_i'')(p_{i+1}p+q_{i+1})\times$$
$$\ldots\times(p_{k-1}p+q_{k-1})p , \qquad (2.3.3)$$

where $p_i+q_i = 1$, $(i = 1,\ldots,k)$.

Hence, $\Pi_{\underline{T}} - \Pi_{\underline{T}'} = [(p_ip+q_i)-(p_i'p+q_i')(p_i''p+q_i'')]X$,

where $X = (p_1p+q_1)\ldots(p_{i-1}p+q_{i-1})(p_{i+1}p+q_{i+1})\ldots(p_{k-1}p+q_{k-1})p > 0$.

Let x = number of players remaining in the i^{th} round $(x > 2)$. Then,

$$p_i = \frac{2m_i}{x}; \quad p_i' = \frac{2m_i'}{x} ; \quad p_i'' = \frac{2m_i''}{x-m_i'} .$$

Using $m_i = m_i' + m_i''$, after simple algebraic manipulations, we have,

$$\Pi_{\underline{T}}-\Pi_{\underline{T}'} = \frac{(1-p)(m_i-m_i')2m_i'(2p-1)}{x(x-m_i')} \cdot X . \qquad (2.3.4)$$

Since $x \geq 2m_i'$; $2p > 1$; $p < 1$; $m_i > m_i'$; $X > 0$, we have $\Pi_{\underline{T}} > \Pi_{\underline{T}'}$, which completes the proof.

<u>Corollary 2.3.2</u>

(1) $\Pi_{\underline{T}'} > \Pi_{\underline{T}}$, for fixed N and p $(0 < p < .5)$; $(N \geq 4)$. (2.3.5)

(2) $\quad \Pi_{\underline{T}'} = \Pi_{\underline{T}}$, for fixed N and p $(p = 0,\frac{1}{2},1;\ N \geq 4)$.

$$(2.3.6)$$

Proof

(1) Exactly as Theorem 2.3.1 with $0 < p < .5$.

(2) When $\underline{p = 0}$, $\Pi_{\underline{T}'} = \Pi_{\underline{T}} = 0$ by (2.3.4).

$\underline{p = 1}$, $\Pi_{\underline{T}'} = \Pi_{\underline{T}} = 1$ by (2.3.2) and (2.3.3) .

$\underline{p = \frac{1}{2}}$, All players are equally likely to win the tournament with probability $\frac{1}{N}$, hence,

$$\Pi_{\underline{T}'} = \Pi_{\underline{T}} = \frac{1}{N} \quad . \tag{2.3.7}$$

(It is clear that if $\underline{T}$, $\underline{T}'$ are any two tournament vectors satisfying Definition 2.1.1 for a fixed N , then (2) still holds when $p = 0,\frac{1}{2},1$.)

From now on we shall only consider the case where $p \in (0.5,1)$, since results for $p \in (0,0.5)$ are obtained in a similar manner. The programs in the Appendix, where applicable, are valid for all values of $p \in [0,1]$.

The descendant principle provides a convenient comparison of tournaments, its main advantage being the absence of calculation. This concludes the preliminary theoretical results useful in the comparison of "tournament patterns".

CHAPTER III

COMPARISON THEOREM AND TWO-PLAYER MODEL

In this chapter we shall describe several examples of "tournament patterns" which have been studied in the literature (e.g. Narayana and Zidek [6]) and are of special importance in the applications. We then state and prove our main theorem which gives a comparison of these "tournament patterns" when one strong player is present. We consider, in the final section, two strong players each with a probability p of defeating an opponent, and probability $\frac{1}{2}$ of defeating one another - the two-player model. Because of the unwieldy formulae involved no attempt is made to compare "tournament patterns" in general in the latter model.

§3.1 Four Special "Tournament Patterns"

A "tournament pattern" is a rule for playing a random knock-out tournament which applies whatever the number of players present. The "tournament patterns" defined below are compared in the next section.

For any $N = 2^t + K \quad (0 \leq K < 2^t;\ t = 1,2,3,\ldots)$,

(1) T_1 is defined to be the tournament with each element m_i of the vector $\underline{m}$ to be $[\frac{N + 2^{i-1} - 1}{2^i}]$ $(i = 1,\ldots,k)$, i.e. taking the maximum number of players possible in each

round (or minimum byes in each round).

Examples when $N = 6,7,8$ are (3,1,1), (3,2,1), (4,2,1) respectively.

(2) T_2 is defined to be the tournament with vector $\underline{m}=(K,2^{t-1},2^{t-2},\ldots,1)$. This tournament may be considered to be a "quick reduction to the classical case" where $N = 2^t$ players - a "tournament pattern" used extensively in such knock-out tournaments as a tennis tournament.

Examples for $N = 5,6,7$ are (1,2,1), (2,2,1) and (3,2,1) respectively.

(3) T_3 is defined to be the tournament played according to the vector $\underline{m} = (1,1,\ldots,1)$, i.e. giving a maximum number of byes in each round.

(4) T_4 is defined to be the tournament played according to the vector $(1,1,\ldots,1)$, with the further restriction that the winner in any round automatically plays in the succeeding round. Thus, after a pair drawn at random plays in the first round, one of the remaining $(N-2)$ bye players of round one meets the winner in round two, and so on. This tournament is not included in Definition 2.1.1.

A comparison of all tournaments satisfying Definition 2.1.1 along with tournament T_4 is possible numerically, for the one strong player model with a fixed $N > 4$ and $p \in [0,1]$, by using programs 1, 2 and 3 of the Appendix. However the ordering of tournaments, using the comparison

criteria of the Introduction, changes for different values of p for some values of N. For example, when $p = .55$, $N = 8$ the tournament vectors (1,2,1,2,1), (3,2,1,1) are ranked 9^{th} and 11^{th} respectively, and when $p = .95$, $N = 8$ they are ranked 10^{th} and 9^{th} respectively (we might say that a "reversal" has occurred). In the Appendix (Table 1) we have provided an example of a table of results for varying values of $p \in (0,1)$ when $N = 7$ in both one- and two-strong player models in the tournament T_4 and those satisfying Definition 2.1.1. From this table a comparison of tournaments may be made for a fixed p when $N = 7$.

It is possible to rank "tournament patterns" in the one-player model in general for $p \in (0.5,1)$, by means of the theorem of the following section.

§3.2 Main Theorem

Theorem 3.2.1

If $\Pi_i(N)$ is the probability that player A wins the tournament T_i, $i = 1,2,3,4$ for all fixed $N \geq 3$, then

$$\Pi_2(N) \geq \Pi_1(N) > \Pi_3(N) > \Pi_4(N) , \qquad (3.2.1)$$

for a fixed p lying in the closed interval $(0.5,1)$.

Note: If $N = 2^M(2^t-1)$, then $\Pi_2(N) = \Pi_1(N)$, $(M = 1,2,\ldots ;$ $t = 1,2,\ldots)$.

Proof

(i) $\underline{\Pi_3(N) > \Pi_4(N)}$ (proof by induction on N)

We note that $\Pi_3(N) = \Pi_4(N)$ when $N = 2,3$. It is true for $N = 4$, as follows:

$$\Pi_3(4) = (\tfrac{2}{4} p + \tfrac{2}{4})(\tfrac{2}{3} p + \tfrac{1}{3})p = \frac{p^3}{3} + \frac{p^2}{2} + \frac{p}{6} .$$

$$\Pi_4(4) = (\tfrac{2}{4} p^3 + \tfrac{2}{4}(\tfrac{1}{2} p^2 + \tfrac{1}{2} p)) = \frac{p^3}{2} + \frac{p^2}{4} + \frac{p}{4} .$$

$$\Pi_3(4)-\Pi_4(4) = p(\frac{-p^2}{6} + \frac{p}{4} - \frac{p}{12}) = \frac{1}{12} p(1-p)(2p-1) > 0 .$$

Similarly we can prove (i) true for $N = 5$.

We assume (i) true for (N-1) players, i.e. $\Pi_3(N-1) > \Pi_4(N-1)$.

$$\Pi_3(N) = (\tfrac{2}{N} p + \tfrac{N-2}{N})\Pi_3(N-1) .$$

$$\begin{aligned} \Pi_4(N) &= (\tfrac{2}{N} p^{N-1} + \tfrac{N-2}{N}(\tfrac{p^{N-2}}{N-2} + \ldots)) , \\ &= \frac{2}{N} p^{N-1} + \frac{p^{N-2}}{N} + \frac{p^{N-3}}{N} + \ldots + \frac{p^2}{N} + \frac{p}{N} , \qquad (3.2.2) \\ &= \frac{1}{N} [(2p-1)p^{N-2} + (N-1)\Pi_4(N-1)] . \end{aligned}$$

When $\Pi_3(N) > \Pi_4(N)$, then

$$(\tfrac{2}{N} p + \tfrac{N-2}{N})\Pi_3(N-1) > \tfrac{1}{N}[(2p-1)p^{N-2} + (N-1)\Pi_4(N-1)],$$

hence,

$$(2p-1)[\Pi_3(N-1)-p^{N-2}]+(N-1)[\Pi_3(N-1)-\Pi_4(N-1)] > 0 ,$$

Proof

(i) $\Pi_3(N) > \Pi_4(N)$ (proof by induction on N)

We note that $\Pi_3(N) = \Pi_4(N)$ when $N = 2,3$. It is true for $N = 4$, as follows:

$$\Pi_3(4) = (\tfrac{2}{4}\, p + \tfrac{2}{4})(\tfrac{2}{3}\, p + \tfrac{1}{3})p = \frac{p^3}{3} + \frac{p^2}{2} + \frac{p}{6} .$$

$$\Pi_4(4) = (\tfrac{2}{4}\, p^3 + \tfrac{2}{4}(\tfrac{1}{2}\, p^2 + \tfrac{1}{2}\, p)) = \frac{p^3}{2} + \frac{p^2}{4} + \frac{p}{4} .$$

$$\Pi_3(4)-\Pi_4(4) = p(\frac{-p^2}{6} + \frac{p}{4} - \frac{p}{12}) = \frac{1}{12}\, p(1-p)(2p-1) > 0 .$$

Similarly we can prove (i) true for $N = 5$.

We assume (i) true for $(N-1)$ players, i.e. $\Pi_3(N-1) > \Pi_4(N-1)$.

$$\Pi_3(N) = (\tfrac{2}{N}\, p + \tfrac{N-2}{N})\Pi_3(N-1) .$$

$$\Pi_4(N) = (\tfrac{2}{N}\, p^{N-1} + \tfrac{N-2}{N}(\tfrac{p^{N-2}}{N-2} + \ldots)) ,$$

$$= \frac{2}{N}\, p^{N-1} + \frac{p^{N-2}}{N} + \frac{p^{N-3}}{N} + \ldots + \frac{p^2}{N} + \frac{p}{N} , \qquad (3.2.2)$$

$$= \frac{1}{N}\, [(2p-1)p^{N-2} + (N-1)\Pi_4(N-1)] .$$

When $\Pi_3(N) > \Pi_4(N)$, then

$$(\tfrac{2}{N}\, p + \tfrac{N-2}{N})\Pi_3(N-1) > \tfrac{1}{N}[(2p-1)p^{N-2} + (N-1)\Pi_4(N-1)],$$

hence,

$$(2p-1)[\Pi_3(N-1)-p^{N-2}]+(N-1)[\Pi_3(N-1)-\Pi_4(N-1)] > 0 ,$$

since,

$$(2p-1) > 0; \quad (N-1) > 0; \quad \Pi_3(N-1) > \Pi_4(N-1);$$

and

$$\Pi_3(N-1) = \left(\frac{2}{N-1}\,p + \frac{N-3}{N-1}\right)\left(\frac{2}{N-2}\,p + \frac{N-4}{N-2}\right)\ \cdots\ \left(\frac{2}{3}\,p + \frac{1}{3}\right)p\ ,$$
$$> p^{N-2}\ ,$$

hence, $\Pi_3(N) > \Pi_4(N)$.

(ii) $\underline{\Pi_3(N) < \Pi_1(N)\ ;\ \Pi_3(N) < \Pi_2(N)}$

The tournament T_3 is a descendant of every vector with the same number of players, N . By Theorem (2.3.1), $\Pi_3(N)$ is less than A's probability of winning in any other tournament with vector $(m_1,\ldots,m_k)$ for the same N , except of course, itself. In particular $\Pi_3(N) < \Pi_1(N)$, and $\Pi_3(N) < \Pi_2(N)$.

(iii) $\underline{\Pi_2(N) \geq \Pi_1(N)}$

At first we shall prove that in tournament T_2 , A has a probability of winning a tournament of N players, which is as great as, and in most cases greater than, any tournament with vector $(m_1,\ldots,m_k)$. (We could then say that T_2 is the "best" tournament when one strong player is present.) Then (iii) follows as a special case of this result.

We shall consider a set of tournaments of the "type" T_2 , i.e. those whose vector $\underline{m}$ is of the form:

(1) $\underline{m} = (x, y, 2^{t-1}, 2^{t-2}, \ldots, 1)$, where $1 \leq x < K$ and
$y = N - x - 2^t = K - x$,

and

(2) $\underline{m} = (x, y, 2^{t-2}, 2^{t-3}, \ldots, 1)$, where $\frac{N}{2} > x > K$ and
$y = N - x - 2^{t-1}$,

then show that T_2 is "better" than (1) or (2), and by induction that T_2 is "better" than the complete class of tournaments of Definition 2.1.1, with the same number of players.

Let $\underset{x<K}{\Pi_2}$ denote the probability that A wins the tournament with vector (1), and let $\underset{x>K}{\Pi_2}$ denote the probability that A wins the tournament with vector (2). Tournaments of the form (1) are first-stage descendants of T_2 , therefore by Theorem (2.3.1),

$$\Pi_2(N) > \underset{x<K}{\Pi_2} \tag{3.2.3}$$

We now show that $\Pi_2(N) > \underset{x>K}{\Pi_2}$,

$$\Pi_2(N) = p^t[\frac{2K}{N}p + \frac{N-2K}{N}] = p^t \frac{2K}{N}(p-1) + p^t \quad . \tag{3.2.4}$$

$$\underset{x>K}{\Pi_2} = p^{t-1}[\frac{2x}{N}p + \frac{N-2x}{N}][\frac{2(N-x-2^{t-1})}{(N-x)}p + 1 - \frac{2(N-x-2^{t-1})}{(N-x)}],$$

$$= \frac{p^{t-1}(p-1)}{N(N-x)}[(N+K-2x)2xp-(N-2x)(x-K)] + p^t \quad . \tag{3.2.5}$$

When $\Pi_2(N) > \underset{x>K}{\Pi_2}$

then, $2Kp < \frac{1}{(N-x)}[(N+K-2x)2xp-(N-2x)(x-K)]$, since $(p-1) < 0$, and,

$$0 < 2p(x-K)(N-x)-2xp(x-K)-(x-K)(N-2x) ,$$

$$\therefore \ 0 < (2p-1)(N-2x) , \quad \text{since} \quad x > K .$$

$2p > 1$, $N > 2x$, hence,

$$\Pi_2(N) > \underset{x > K}{\Pi_2} , \quad (x < \tfrac{N}{2}) . \tag{3.2.6}$$

We note that when $x = \frac{N}{2}$, and using (3.2.5),

$$\underset{x > K}{\Pi_2} = 2p^{t-1} \frac{(p-1)}{N^2} NpK + p^t , \tag{3.2.7}$$

$$= \frac{2K}{N} p^t(p-1)+p^t = \Pi_2(N) \quad \text{by (3.2.4)} .$$

We now see that for any allowable x , followed by a "quick reduction to the classical case", T_2 is the "best" tournament. We shall show that T_2 is the "best" tournament in general by induction on N .

When $N = 4$, we have two tournament vectors (2,1) and (1,1,1); by (3.2.3), T_2 is "best". When $N = 5$, we have three tournament vectors (2,1,1), (1,2,1) and (1,1,1,1). Of the vectors whose first component is 2 , (2,1,1) is the "best", and of the vectors whose first component is 1 , T_2 is "best" by deleting the first component and considering the resulting tournament vectors for $N = 4$.

We now assume that T_2 is "best" for $5,6,\ldots,N-1$ players. If we let T_2' denote the class of tournaments of the form (1), (2), T_2 or those with $x = \frac{N}{2}$ in (2) for a fixed N, then $\Pi_2(N) \geq \Pi_2'(N)$ where the notation Π_2' is obvious. If we reduce each tournament vector of T_2' by its first element, we shall be left with tournaments of the pattern T_2, or at least a tournament where the probability that the strong player wins is equivalent to that in T_2. Then, by the induction hypothesis, any member of the class of tournaments T_2' is better than any other tournament with the same first element satisfying Definition 2.1.1; since if we reduce any particular tournament vector not in T_2' by its first element to a tournament of $N'(< N)$ players, it will be dominated by a tournament in T_2' which has been reduced by the same first element.

$$\therefore \quad \Pi_2(N) \geq \Pi_{(m_1,m_2,\ldots,m_k)}(N) ,$$

and in particular,

$$\Pi_2(N) \geq \Pi_1(N) \quad .$$

(iv) $\underline{\Pi_1(N) = \Pi_2(N) \text{ if } N = 2^M(2^t-1)}$

Let m_1', m_1'' be the first components in the tournament vectors describing T_1, T_2 respectively. It then follows that if $m_1' = m_1''$, T_1 and T_2 have the same tournament vector.

If $N = 2^t$ $(t = 1,2,\ldots)$, $m_1' = m_1'' = 2^{t-1}$, by 3.1(1) and 3.1(2).

If $N = 2^t-1$ $(t = 2,3,\ldots)$, from 3.1(1), we have for T_1,

$$m_1' = [\frac{N}{2}] = [\frac{2^t-1}{2}] \equiv (\frac{1}{2}\cdot 2^t)-1 = 2^{t-1}-1 \quad ,$$

and from 3.1(2), we have for T_2 ,

$$m_1'' = K = (2^t-2^{t-1})-1 = 2^{t-1}-1 \quad .$$

Hence,

$$\Pi_1(N) = \Pi_2(N) \quad \text{for} \quad N = 2^t, 2^t-1 \quad . \qquad (3.2.8)$$

By (3.2.7) and (3.2.8), when $N = 2^M(2^t-1)$, $(t = 1,2,\ldots;$ $M = 1,2,\ldots)$,

$$\Pi_1(N) = \Pi_2(N) \quad .$$

This completes the proof of the theorem.

Initial computer results (up to $N = 256$) suggest that $\Pi_2(N) > \Pi_1(N)$ for $N \neq 2^M(2^t-1)$, but a proof is still awaited.

If $E(R_i)$, $(i = 1,2,3,4)$, is the expected number of games played by the strong player in the course of the tournament T_i , using Theorems (2.1.2) and (3.2.1) we have $E(R_4) > E(R_3) > E(R_1) \geq E(R_2)$ for $p \in (0.5,1)$ and N fixed (Narayana and Zidek [6], p. 14 have stated that Theorem (2.1.2) is true for tournament T_4). This result is of importance in the design of knock-out tournaments when the criteria for choosing the best player depends on the number of games won, thus a knock-out tournament or a repeated knock-out tournament of T_4 would increase the expectation of the best player as opposed to T_3, T_1 or T_2 .

A natural extension of the previous theorem would be to ask whether T_2 is the "best" tournament of all random knock-out tournaments or, similarly, whether T_4 is the "worst" tournament from the point of the strong player's probability of winning , (in the one-player model).

§3.3 <u>The Two-Player Model</u>

We have discussed the probability that a strong player wins a tournament with vector $(m_1,...,m_k)$ as well as tournament T_4 when there is one strong player present. It could happen that two players of equal strength are present, each having a probability p of defeating an opponent - the 'two-player' model. We shall consider in the latter model the probability $\Pi^{**}_{(m_1,...,m_k)}$ that one of the strong players wins a tournament with vector $(m_1,...,m_k)$ and give explicit results for the same probability in the classical tournament where $N = 2^t$ players, and in tournament T_4 . We hope that these probabilities will be useful in studying the design of repeated random knock-out tournaments (possibly by simulation using random numbers), when there are two strong players present.

(a) <u>A Recursive Formula for $\Pi^{**}_{(m_1,...,m_k)}$</u>

We shall put $\Pi^{*}_{(m_j,...,m_k)}$ = probability of one strong player (A, say) winning the tournament with remaining vector $(m_j,...,m_k)$ $(j = 2,...,k)$, if he is the only strong player present, with similar meaning for $\Pi^{**}_{(m_j,...,m_k)}$.

If there are N players $A, B, C_1, C_2, \ldots, C_{N-2}$; with A and B having probability p of defeating an opponent, and $\frac{1}{2}$ of defeating one another, using the results by Capell and Narayana [1] that:

(i) Both A, B play in round one with probability $\binom{N-2}{2m_1-2}/\binom{N}{2m_1}$,

(ii) Exactly one of A, B plays in round one with probability $2\binom{N-2}{2m_1-1}/\binom{N}{2m_1}$,

(iii) Neither A nor B plays in round one with probability $\binom{N-2}{2m_1}/\binom{N}{2m_1}$,

then a recursive formula for $\Pi^{**}_{(m_1,\ldots,m_k)}$ may be found if we consider the following mutually exclusive events in round one:

(1) A and B both play, they both meet then A wins, or they do not meet and B wins or loses,

(2) A or B plays and wins,

(3) Only B plays and loses,

(4) Neither A nor B plays,

Hence,

$$\Pi^{**}_{(m_1,\ldots,m_k)} = [\binom{N-2}{2m_1-2}/\binom{N}{2m_1}]\{\frac{1}{2m_1-1}\cdot\frac{1}{2}\cdot\Pi^{*}_{(m_2,\ldots,m_k)} + \frac{2m_1-2}{2m_1-1}\,p[p\Pi^{**}_{(m_2,\ldots,m_k)}+q\Pi^{*}_{(m_2,\ldots,m_k)}]\} +$$
$$+[2\binom{N-2}{2m_1-1}/\binom{N}{2m_1}]\{p\Pi^{**}_{(m_2,\ldots,m_k)}\} + [\binom{N-2}{2m_1-1}/\binom{N}{2m_1}]q\Pi^{*}_{(m_2,\ldots,m_k)} +$$
$$+ [\binom{N-2}{2m_1}/\binom{N}{2m_1}]\,\Pi^{**}_{(m_2,\ldots,m_k)}, \text{ where } q = 1-p, \qquad (3.3.1)$$

for computation this is expressed as,

$$\Pi^{**}_{(m_1,\ldots,m_k)} = \Pi^{*}_{(m_2,\ldots,m_k)}\{\frac{m_1}{N(N-1)}(1+2qp(2m_1-2)+q\binom{N-2}{2m_1-1}/\binom{N}{2m_1})\}$$

$$+ \Pi^{**}_{(m_2,\ldots,m_k)}\{p^2 \frac{2m_1(2m_1-2)}{N(N-1)}+[2\binom{N-2}{2m_1-1}p+\binom{N-2}{2m_1}]\div\binom{N}{2m_1}\}. \quad (3.3.2)$$

A program for (3.3.2) is given in the Appendix (Program 6).

(b) Classical Tournament ($N = 2^t$ Players, $t = 1,2,\ldots$)

We shall consider two methods,

(i) Recursive Method

Let $\Pi^*(t)$ = probability that a strong player wins the classical tournament if there is only one strong player present (all other players being of equal strength) = p^t, and $\Pi^{**}(t)$ = probability that A wins in the two-player model. Considering the two mutually exclusive events that:

(1) A and B meet,

(2) They do not meet and B wins or loses,

we have,

$$\Pi^{**}(t) = \frac{1}{(2^t-1)} \cdot \frac{1}{2}\,\Pi^*(t-1) + \frac{2^t-2}{2^t-1}\,p[p\Pi^{**}(t-1)+q\Pi^*(t-1)] \,. \quad (3.3.3)$$

(ii) Explicit Method

The probability that A meets B in the k^{th} round ($k = 1,\ldots,t$)

$$= \frac{2^{k-1}\,p^{2k-2}}{2^t-1} \,,$$

therefore the probability that A wins the tournament, given he meets B in the k^{th} round,

$$= \frac{2^{k-1} p^{2k-2}}{2^t - 1} \cdot \frac{p^{t-k}}{2} \quad . \tag{3.3.4}$$

The probability that A does not meet B by the i^{th} round, which B loses,

$$= \left(\frac{2^{t-i}-1}{2^t-1}\right) \circ 2^i \, p^{i-1} q \qquad (i = 1,2,\ldots,t-1) \ ,$$

therefore the probability that A and B do not meet

$$= \sum_{i=1}^{t-1} \frac{(2^{t-i}-1)}{(2^t-1)} \cdot 2^i \, p^{i-1} q \quad . \tag{3.3.5}$$

Hence the probability that A wins the tournament using (3.3.4) and (3.3.5),

$$\Pi^{**}(t) = \sum_{k=1}^{t} \frac{2^{k-1} p^{2k-2}}{2^t-1} \frac{p^{t-k}}{2} + p^t \sum_{i=1}^{t-1} \frac{(2^{t-i}-1)}{(2^t-1)} 2^i p^{i-1} q \ ,$$

which reduces to,

$$\Pi^{**}(t) = \frac{p^{t-1}}{2^{t+1}-2} \frac{(1-(2p)^t)}{1-2p} + \frac{(2p)^t}{2^t-1}(1-p^{t-1}) - \frac{p^t q 2(1-(2p)^{t-1})}{(2^t-1)(1-2p)} \ . \tag{3.3.6}$$

(c) <u>Tournament T_4</u>

If we order the N players in the positions $1,2,\ldots,N$, there will be N! possible arrangements. We shall

let B be in position i (= 1,2,...,N-1) , A in position j (= 2,...,N) such that $i < j$.

If B is in position $i = 1$, he will meet A with probability $p^{j-2} = p^{j-(i+1)}$. If $i > 1$, then player B will meet A with probability p^{j-i} .

Hence if p_{ij} = probability that B meets A , then

$$p_{ij} = \begin{cases} p^{j-(i+1)} & \text{if } i = 1;\ j = 2,\dots,N \,. \\ p^{j-i} & \text{if } i = 2,\dots,N-1\,;\ j = 3,\dots,N;\ \ i < j \,. \end{cases} \qquad (3.3.7)$$

Similarly the probability that B does not meet A is $(1-p_{ij})$, i.e. B is eliminated.

The probabilities $\{p_{ij}\}$ in (3.3.7) may be written as a column vector ($\underline{\alpha}$, say) as follows:

$$\{p_{ij}\} = \underline{\alpha} = \begin{bmatrix} p_{1,2} \\ p_{1,3} \\ p_{1,4} \\ \vdots \\ p_{1,N} \\ p_{2,3} \\ p_{2,4} \\ \vdots \\ p_{2,N} \\ \vdots \\ p_{i,i+1} \\ \vdots \\ p_{i,j} \\ \vdots \\ p_{i,N} \\ p_{i+1,i+2} \\ \vdots \\ p_{i+1,N} \\ \vdots \\ p_{N-1,N} \end{bmatrix} = \begin{bmatrix} p^{o} \\ p^{1} \\ p^{2} \\ \vdots \\ p^{N-2} \\ p^{3-2} \\ p^{4-2} \\ \vdots \\ p^{N-2} \\ \vdots \\ p^{(i+1)-i} \\ \vdots \\ p^{j-i} \\ \vdots \\ p^{N-i} \\ p^{(i+2)-(i+1)} \\ \vdots \\ p^{(i+1)-N} \\ \vdots \\ p^{N-(N-1)} \end{bmatrix}_{[\frac{1}{2}(N-1)(N-2)] \times 1} . \qquad (3.3.8)$$

If B meets A , who is in position j , he defeats A with probability $\frac{1}{2}$, then wins the remaining rounds with probability p^{N-j} . If B is defeated, then A who must be

in position j, wins the tournament with probability $p^{(N-j)+1}$. (If B and A are in positions 1,2 respectively, they meet with probability 1, and do not meet with probability 0.) Again the probabilities $\{p^{N-j}\}$, $\{p^{(N-j)+1}\}$ may be written as the column vectors $\underline{\beta}$ and $\underline{\gamma}$ respectively, where,

$$\{p^{(N-j)}\} = \underline{\beta} = \begin{bmatrix} \left.\begin{matrix} p^{N-2} \\ p^{N-3} \\ \vdots \\ p^{N-N} \end{matrix}\right\} \\ \left.\begin{matrix} p^{N-3} \\ \vdots \\ p^{o} \end{matrix}\right\} \\ \vdots \\ \left.\begin{matrix} p^{N-j} \\ \vdots \\ p^{o} \end{matrix}\right\} \\ \left.\begin{matrix} p^{N-j-1} \\ \vdots \\ p^{o} \end{matrix}\right\} \\ \vdots \\ \left.p^{o}\right\} \end{bmatrix}_{[\frac{1}{2}(N-1)(N-2)] \times 1} . \qquad (3.3.9)$$

Clearly, $$\underline{\gamma} = p\,\underline{\beta} . \qquad (3.3.10)$$

Since B could be in A's position, the probability

that the two strong players meet, and that A wins the tournament given that they met,

$$= \frac{2}{N!} p_{ij} \cdot \frac{1}{2} \cdot p^{N-j} \qquad \begin{matrix}(i = 1,\ldots,N-1); \\ (j = 2,\ldots,N);\end{matrix} \qquad (i < j) .$$

If they do not meet, the probability that A wins the tournament,

$$= \frac{1}{N!} (1-p_{ij}) \, p^{(N-j)+1} , \qquad \text{(i.e. } B \text{ is at position } i \text{ , and } A \text{ at } j) .$$

Therefore the probability $\Pi_4^{**}(N)$, that A wins the tournament T_4,

$$= \frac{1}{N!} \{ 2\underline{\alpha}'\underline{\beta} \cdot \frac{1}{2} + (\underline{1}-\underline{\alpha})'\underline{\gamma} \} ,$$

$$\text{where } \underline{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}^{[\frac{1}{2}(N-1)(N-2)] \times 1} .$$

Using (3.3.10), we have,

$$\Pi_4^{**}(N) = \frac{1}{N!}(p\underline{1} + q\underline{\alpha})'\underline{\beta} , \qquad \text{where } q = 1-p .$$

Since the (N-2) "weak" players may be arranged in (N-2)! ways, we have,

$$\Pi_4^{**}(N) = \frac{1}{N(N-1)} \{(p\underline{1} + q\underline{\alpha})'\underline{\beta}\} . \tag{3.3.11}$$

which, using (3.3.8) and (3.3.9), becomes,

$$\Pi_4^{**}(N) = \frac{1}{N(N-1)} \{p \sum_{\ell=o}^{N-2} \sum_{s=o}^{\ell} p^s + q(N-1)p^{N-2} + q \sum_{m=2}^{N-1} (N-m)p^{N-m}\} \quad , \tag{3.3.12}$$

which reduces to,

$$\Pi_4^{**}(N) = \frac{1}{N(N-1)} \{\frac{p}{q} [(N-1) - \frac{p}{q} (1-p^{N-1})] + + q(N-1)p^{N-2} + [\frac{p(1-p^{N-2})}{1-p} - (N-2)p^{N-1}]\}. \tag{3.3.13}$$

A computer program is given in the Appendix for (3.3.13), (Program 7).

CHAPTER IV

SIMULATION OF REPEATED TOURNAMENTS

In this chapter we hope to give some insight into the design of tournaments by simulating them using random digits. Narayana and Zidek [7] have suggested that a desirable design of a tournament tends to increase the expected number of plays involving the best player. With this basic idea in mind we shall describe three tournaments and compare the effectiveness of them to a round-robin tournament (where every player meets every other player once). We shall use two criteria to compare them which have been developed by David [2], one following Bechhofer the other following Gupta, for a round-robin tournament. Firstly we shall count the number of times (x) a strong player has the highest score, and secondly the number of players (y) included in a "best" subset S after a series of simulations of these tournaments.

§4.1 Selection Rules

When our aim is the selection of the "best" object (or strongest player in tournaments) with at least a predetermined probability P' , we shall use tables presented by David giving the smallest number of replications (n) of a round-robin tournament with t players. In order to compare

the repeated knock-out tournaments when this selection rule is used, we shall play the same number of matches as the round-robin with the same number of players.

We are now interested in finding the best subset S, which is just large enough to ensure, with at least a pre-assigned probability P^*, that the best object $A_{(t)}$ is included in S. Following David, giving a score of $+1$ to the winner of one comparison, 0 to the other, we use the decision rule:

Retain in S only those objects A_i for which their score $a_i \geq a_{(t)} - \nu'$, where $a_{(t)}$ is the highest score and ν', a non-negative integer, is a function of t, n and P^*.

Tables are available in David for the value ν', for a range of values for t, n and P^*.

If the win 1, lose 0 scoring system used in round-robin tournaments is applied to knock-out tournaments, a player who plays in nine matches and loses eight of them for example, would be ranked equal with a player who had played and won his only match. To overcome difficulties such as these we shall give a score of $+1$ for a win and -1 for a defeat in a single match, ν' then being replaced by $\nu = 2\nu'$.

§4.2 <u>Simulation Using a Table of Random Digits</u>

In this section we shall investigate four "tournament

patterns" for their effectiveness in selecting the best player out of three and then four. In agreement with the basic model of David we shall consider one strong player who has a probability p of defeating an opponent, all other players being of equal strength. For one complete simulation we shall consider the two strong player model of §3.3 .

We shall study the effectiveness of repeated trials of the round-robin tournament (RR), and the knock-out tournaments T_2 and T_4 of §3.1 with the exception that the winner of each completed tournament of T_4 automatically plays in the first round of the next repetition (T_{4K}) . Finally we shall consider a "deterministic" tournament (D), for the first repetition of which we play T_4 , the winner then plays against one of the remaining players who has the highest score. We carry on in this way playing the winner of the previous round against the player with the highest score until as many matches as in the round-robin have been played (similarly for T_2 and T_{4K}) . If, in the "deterministic" tournament we are in the situation where we have a winner of the previous round and more than one other player with the same score, we choose a player at random.

It seems reasonable that T_2 and a slightly modified T_4 should be compared when the criteria for selecting a player is dependent on the score, since in general $E(R_4) > E(R_2)$ for $p \in (0.5,1)$ by the theorem of §3.2. Also a natural extension to T_{4K} is the "deterministic"

tournament where $E(R)$ is comparatively large and a player who does well in the early stages of the tournament plays as often as possible, hence giving a possible strong player a good chance of coming to the forefront and playing more matches, in compliance with the suggestion of Narayana and Zidek. We examined further the effectiveness of these tournaments when the model was extended to two strong players, since in many selection problems more than one object may be considered as being equally good.

Random digits, from a Book of Random Numbers [8], were used to provide the probability of one player defeating another, and also for selecting players at random to play in the knock-out tournaments. As far as possible the same digits were used for each tournament simulation. Ten different simulations of all four tournaments were made in order to make a reasonable comparison.

Example

A simulation for four players (A,B,C,D), for the tournaments RR, T_2, T_{4K} and D. A is the strong player with $p = .75$. From the tables of David with $P' = .90$, $P^* = .75$ we have $n = 6$ and $\nu = 8$.

Notation: $A \to B$ shall mean that A defeats B .

Round-Robin Simulation

1.	A → B	A → C	B → C	A → D	D → B	C → D
2.	A → B	A → C	B → C	A → D	D → B	C → D
3.	A → B	A → C	C → B	D → A	D → B	D → C
4.	B → A	A → C	C → B	A → D	B → D	C → D
5.	A → B	A → C	C → B	A → D	B → D	D → C
6.	A → B	C → A	C → B	A → D	D → B	D → C

Players	A	B	C	D
Scores	12	-8	-2	-2

$x = 1\ ,\quad y = 1\ .$

T_2 Simulation

	Round I		Round II	Winner
1.	C → D	A → B	A → C	A
2.	D → B	A → C	A → D	A
3.	C → D	A → B	A → C	A
4.	A → D	B → C	A → B	A
5.	D → A	B → C	D → B	D
6.	C → B	A → D	A → C	A
7.	D → C	B → A	D → B	D
8.	A → B	C → D	A → C	A
9.	A → B	D → C	A → D	A
10.	A → B	D → C	A → D	A
11.	B → D	C → A	C → B	C
12.	B → A	D → C	B → D	B

Players	A	B	C	D
Scores	12	-5	-5	-2

$x = 1\ ,\quad y = 1\ .$

T_{4K} Simulation

	Round I	Round II	Round III	Winner
1.	C → D	C → B	A → C	A
2.	A → C	A → D	A → B	A
3.	A → D	A → B	A → C	A
4.	D → A	C → D	B → C	B
5.	B → A	D → B	D → C	D
6.	D → B	A → D	A → C	A
7.	A → C	A → B	A → D	A
8.	A → D	A → C	A → B	A
9.	A → B	D → A	C → D	C
10.	C → A	C → B	D → C	D
11.	D → B	D → C	D → A	D
12.	A → D	B → A	C → B	C

Players	A	B	C	D
Scores	11	-8	-3	0

$x = 1$, $y = 1$.

D Simulation

	Round I	Round II	Round III	Winner
1.	C → D	B → C	A → B	A
2.	A → C	A → B	A → D	A
3.	A → B	A → C	A → D	A
4.	A → C	B → A	D → B	D
5.	D → A	D → B	C → D	C
6.	A → C	A → D	A → B	A
7.	A → D	A → C	A → B	A
8.	A → D	C → A	B → C	B
9.	A → B	D → A	D → C	D
10.	D → A	D → B	C → D	C
11.	C → A	D → C	B → D	B
12.	B → A	B → D	C → B	C

Players	A	B	C	D
Scores	9	-4	-3	-2

$x = 1$, $y = 1$.

In the following table $\bar{\nu} = \frac{y}{10}$ is the average number of players included in the best subset S , x is the number of times a strong player had the highest score, and X is the percentage that a strong player was included in the best subset (only given for $t = 3$, $p = .55$ since all others were 100% correct). When $p = .55$, $t = 3$ the value of n for $P' = .90$ was 165 , hence only $\bar{\nu}$ is included in the table for this case.

Table 4.2.1

Numerical Comparison of RR, T_2, T_{4K} and D , $(P^* = .75)$.

	t	p	n	ν	RR	T_2	T_{4K}	D
	3	.55	(6)	6	$\bar{\nu}$ = 2.1	2.1	2.1	2.1
					X = 90%	80%	80%	90%
P' = .90	3	.75	6	6	$\bar{\nu}$ = 1.5	1.2	1.3	1.1
					x = 9.5	10	9.5	10
P' = .90	4	.75	6	8	$\bar{\nu}$ = 1.7	1.3	1.3	1.1
					x = 10	9	10	10
P' = .90	4	.75	6	8	$\bar{\nu}$ = 2.1	2.1	2.0	1.9
	(2 strong players)				x = 10	10	10	10

Six repetitions of the tournaments when $t = 3$ and $p = .55$, although far less than recommended by David, showed remarkable consistency in giving the same player the highest score in each tournament of a simulation, throughout the ten

simulations (the strong player being chosen 70% of the time). Hence, in perhaps the least favorable case, all the designs considered seem to be equally desirable. Again, when our object is the selection of the best player with a pre-assigned probability $P' = .90$, little comparison may be made when $t = 4$, $p = .75$ in both models since each of the four tournaments gives the highest score to the strong player not less than 90% of the time.

When our aim is the selection of the best subset with at least a predetermined probability P^* , the results indicate that the "deterministic" rule is perhaps more effective. In the two-player model it is probably more desirable to choose at least one of the strong players, all the tournaments either doing this or including exactly one in the subset. The subset chosen, on average, is significantly higher than that for the one strong player model when $t = 4$, $p = .75$, thus a possible decision rule could be to assume a two-player model if $\bar{\nu} \geq 1.9$, if the model is not known beforehand. Overall, it seems that there is little to choose between the three knock-out tournaments, although the results indicate that they are more effective than the round-robin in selecting a smaller subset containing the strong player, with, perhaps, D being slightly superior.

§4.3 <u>Simulation by Computer</u>

In this section we shall present a table of comparison between the tournaments RR and T_{4K} in the one strong player

model, when we require a subset including the best player with probability $P^* = .75$. This table, calculated in conjunction with H. Morin, was prepared using random digits generated by a computer.

Table 4.3.1

Comparison of RR and T_{4K} $(P' = P^* = .75)$

		p = .55		.60		.65		.70		.75	
		n=71	ν=28	18	14	8	10	5	8	3	6
t = 4	RR	$\bar{\nu}$ = 2.7		2.4		2.4		2.5		2.6	
		142	28	36	14	16	10	10	8	6	6
	T_{4K}	2.6		1.9		2.6 *		2.1		2.2 *	
t = 5		68	33	17	18	8	12	4	8	3	8
	RR	2.2		3.3		2.7		3.3		3.9	
		120	33	43	18	20	12	10	8	8	8
	T_{4K}	2.7		2.4		2.8		1.8		2.5	
t = 6		65	39	16	20	7	12	4	10	3	8
	RR	3.1		3.0		3.5		4.1		3.3	
		195	39	48	20	21	12	12	10	9	8
	T_{4K}	3.4		3.2		2.8		3.5		1.9	
t = 7		61	42	15	22	7	14	4	10	3	10
	RR	3.3		4.0		3.6		2.3		4.2	
		214	42	53	22	25	14	14	10	11	10
	T_{4K}	3.4		3.7		3.3		4.3		2.2	
t = 8		58	45	15	24	7	16	4	12	3	10
	RR	4.4		4.1		5.2		5.4		4.4	
		232	45	60	24	28	16	16	12	12	10
	T_{4K}	5.3		3.4		2.7		2.9		3.3	

* - best player not included in best subset in one out of ten simulations.

Ten simulations were, again, done in each case.

From the table it seems that T_{4K} is more effective in selecting a smaller subset than RR , selecting a smaller subset size 17 times, out of 25 simulations. However both tournaments included the strong player in the best subset over 99% of the time, indicating that David's table ([2], p. 115), which gives $P^* = .75$, is extremely conservative.

APPENDIX

PROGRAMS AND NUMERICAL EXAMPLES

Program 1.

The program *VECTOR* generates all the tournament vectors $(m_1,...,m_k)$ for $N > 4$, which satisfy Definition 2.1.1.

```
     ∇ VECTOR N;E;M;S;V;V1;X
[1]    V←(S←⍳N-2)[1],⍳0
[2]    E←1+V[⍴V]
[3]    →(~(E∊S))/12
[4]    M←V1-0,X[1+⍳¯1+⍴X←¯1⌽V1←V←V,E]
[5]    →(~(∧/(2×M)≤1+V1))/11
[6]    M←V1-0,X[1+⍳¯1+⍴X←¯1⌽V1←V1,N-1]
[7]    →(~(∧/(2×M)≤1+V1))/2
[8]    →((N-1)≠+/M)/2
[9]    ⌽M
[10]   →2
[11]   →3,V←(~(V=¯1+E←1+V[⍴V]))/V
[12]   →(1<⍴V)/11
     ∇
```

<u>Example</u>. N = 7 .

VECTOR 7

3	1	1	1		
2	1	1	1	1	
1	1	1	1	1	1
1	2	1	1	1	
2	2	1	1		
1	1	2	1	1	
1	3	1	1		
3	2	1			
2	1	2	1		
1	1	1	2	1	
1	2	2	1		

Program 2.

The program *PROB* gives the probability Π_N that A wins the tournament with vector $(m_1,...,m_k)$ when he is the only strong player present. We use the expression (2.1.4).

```
     ∇ R←P PROB V;I;L;M;N
[1]    →(1=ρV,ι0)/0,R←P,ρN←I+0.5×+/M←0,V×1+I←1
[2]    →(I<ρV)/2,R←R×(1-L)+P×L←M[I←I+1]÷N←N-0.5×M[I]
     ∇
```

Example. $N = 9$; $p = .25, .50, .65, .85, .95$.

Tournaments T_2, T_1, T_3 .

```
      5 RND P PROB 1 4 2 1
0.01302  0.11111  0.25327  0.59365  0.84785
      5 RND P PROB 4 2 1 1
0.01667  0.11111  0.24717  0.58344  0.84242
      5 RND P PROB 1 1 1 1 1 1 1 1
0.02182  0.11111  0.23654  0.56274  0.83055
```

The program *RND* rounds off the output to five decimal places.

Program 3.

The program *T*4 gives the probability Π_4 that A wins the tournament T_4, when there are N players present. We use $\Pi_4 = (p^{N-1} + p - 2p^N)/N(1-p)$, as in (3.2.2).

```
    ∇ R←P T4 N
[1]   R←((P*¯1+N)+P-2×P*N)÷N×1-P
    ∇
```

Example. N = 9; p = .25, .5, .65, .85, .95 .

```
      5 RND P T4 9
0.03704  0.11111  0.20331  0.48834  0.78427
```

Program 4.

The program *BYE* gives the bye vector $(b_o,\dots,b_{k-1})$ of (2.1.8) for any tournament vector $(m_1,\dots,m_k)$.

```
      ∇ V←BYE A;B;C;D;I;S
    [1]   V←(1-C),C←(B-2×A[1])÷B←1++/A,I←0
    [2]   S←(D←(B-2×A[I+1])÷B←B-A[I←I+1]),(¯1+ρV)ρ0
    [3]   →((ρA)>ρV←+/[1](1-ιρV)⌽V∘.×(1-D),S)/2
      ∇
```

Example. N = 9 .

Tournaments T_2, T_1, T_3 .

```
    5 RND BYE 1 4 2 1
0.22222  0.77778  0   0
    5 RND BYE 4 2 1 1
0.47407  0.41481  0.1037   0.00741
    5 RND BYE 1 1 1 1 1 1 1 1
0.00071  0.00988  0.05679  0.17284  0.29846  0.28951  0.14405  0.0277ε
```

Program 5.

The program *RNI* illustrates how the *BYE* program is used to calculate the probabilities R_N^i, that A plays exactly i rounds in the course of a tournament of N players with vector $(m_1,\ldots,m_k)$. Narayana and Zidek [6] have shown that,

$$R_N^i = (p^i\, b_{k-i} + p^{i-1} q c_{k-i}) \quad ,$$

$$\text{where} \quad c_{k-1} = 1,\; c_{k-i} = 1 - \sum_{j=1}^{i-1} b_{k-j} \quad (i \geq 2)\ .$$

```
    ∇ R←P RNI V;B;X
[1]   R←(P*ιρV)×((1-X+.×(ιρX)∘.≤ιρX←0,B[ι¯1+ρB])÷P÷1-P)+B←⌽BYE V
    ∇
```

Example. $N = 9$; $p = .75$.

Tournaments T_2, T_1, T_3 .

```
      5 RND .75 RNI 1 4 2 1
0.25  0.1875  0.46875  0.09375
      5 RND .75 RNI 4 2 1 1
0.25556  0.24444  0.3   0.2
      5 RND .75 RNI 1 1 1 1 1 1 1 1
0.27083  0.26332  0.2386  0.15125  0.06002  0.0141  0.00179 9E¯5
```

<u>Program 6</u>.

The program *TSP* gives the probability Π^{**} that A wins the tournament $(m_1,\ldots,m_k)$ when there are two strong players present, using the formula (3.3.2).

```
    ∇ R←P TSP V;A;C;D;E;H;I;J;N;Q;T;X;Z
[1]   R←V[I←ρV,T←⍳0]÷2
[2]   X←P PROB T←V[I],T
[3]   N←1++/(H←V[I←I-1]),T
[4]   J←R×((P*2)×A×(A-2)÷Z←N×N-1)+((2×P×D←(A-1)!C)+A!C←N-2)÷E←(A←2×H)!N
[5]   →(1<I)/2,R←(X×((A÷2×Z)×1+2×P×Q×A-2)+(Q←1-P)×D÷E)+J
    ∇
```

<u>Example</u>. $N = 9$; $p = .25, .50, .65, .85, .95$.

Tournaments T_2, T_1, T_3 .

```
      5 RND P TSP 1 4 2 1
0.01579  0.11111  0.22192  0.40313  0.47727
      5 RND P TSP 4 2 1 1
0.01942  0.11111  0.217  0.39587  0.47369
      5 RND P TSP 1 1 1 1 1 1 1 1
0.02455  0.11111  0.20878  0.38334  0.46764
```

Program 7.

Program *TT*4 gives the probability Π_4^{**} that A wins the tournament T_4 when there are two strong players present, using the formula (3.3.13).

```
    ∇ R←P TT4 N;A;B;C;Q
[1]   A←(P÷Q)×((N-1)-(P÷Q←1-P)×1-P*N-1)
[2]   B←Q×(N-1)×P*N-2
[3]   C←((P÷Q)×(1-P*N-2))-(N-2)×P*N-1
[4]   R←(A+B+C)÷N×N-1
    ∇
```

Example. N = 9 ; p = .25, .50, .65, .85, .95 .

```
      5 RND P TT4 9
0.04013  0.11111  0.18331  0.33749  0.44252
```

Table 1.

Using Programs 1, 2, 3, 6 and 7 we may obtain a complete numerical comparison of all tournaments with vector $(m_1,\ldots,m_k)$ and tournament T_4 for the one strong player and two strong player models, for a fixed N and values of $p \in [0,1]$. In this table we display the results obtained when these programs are used for $N = 7$ and $p = .25, .55, .65, .75, .85, .95$.

Numerical Comparison of Tournaments When $N = 7$.

Tournament	Model	p=.25	.55	.65	.75	.85	.95
3 1 1 1	1	0.0279	0.18329	0.28779	0.42969	0.61664	0.857
	2	0.03348	0.17526	0.24867	0.32924	0.40895	0.47602
2 1 1 1 1	1	0.03125	0.18175	0.28285	0.42187	0.60815	0.8524
	2	0.03662	0.17402	0.24501	0.32401	0.40391	0.47367
1 1 1 1 1 1	1	0.03223	0.18123	0.28109	0.41895	0.60483	0.85052
	2	0.0376	0.17361	0.24374	0.3221	0.40199	0.47271
1 2 1 1 1	1	0.03069	0.18201	0.28368	0.42318	0.60957	0.85316
	2	0.03607	0.17423	0.24564	0.32494	0.40488	0.47417
2 2 1 1	1	0.02857	0.18304	0.28704	0.42857	0.6155	0.85641
	2	0.03395	0.17506	0.24815	0.32866	0.40865	0.47607
1 1 2 1 1	1	0.02946	0.18251	0.28525	0.4256	0.61213	0.85452
	2	0.03487	0.17464	0.24685	0.32674	0.40671	0.47511
1 3 1 1	1	0.02455	0.18452	0.29152	0.43527	0.62238	0.85995
	2	0.03051	0.17627	0.25157	0.33333	0.41289	0.4779
3 2 1	1	0.02232	0.18582	0.29575	0.44196	0.62961	0.86382
	2	0.0279	0.17731	0.25477	0.33817	0.41791	0.48052
2 1 2 1	1	0.025	0.18427	0.29068	0.43393	0.62094	0.85918
	2	0.03069	0.17606	0.25104	0.33281	0.41275	0.47811
1 1 1 2 1	1	0.02578	0.18373	0.28886	0.43092	0.61754	0.85728
	2	0.03157	0.17564	0.24974	0.33086	0.41078	0.47713
1 2 2 1	1	0.02455	0.18452	0.29152	0.43527	0.62238	0.85995
	2	0.0302	0.17627	0.25168	0.33377	0.41374	0.47862
T_4	1	0.04764	0.17372	0.25607	0.37772	0.55809	0.82405
	2	0.05298	0.1676	0.22529	0.29419	0.37299	0.45769

1, 2 denote the one- and two-strong player models, respectively. (When $p = .25$, we have a weak player present.)

BIBLIOGRAPHY

[1] Capell, P. and Narayana, T.V. "Contributions to the Theory of Knock-Out Tournaments". Unpublished, The University of Alberta, 1968.

[2] David, H.A. "The Method of Paired Comparisons". London, Charles Griffin and Company Ltd., 1963.

[3] Feller, W. "An Introduction to Probability Theory and its Applications". Volume I, Second Edition, New York, John Wiley and Sons, Inc., 1957.

[4] Iverson, K.E. "A Programming Language". New York, John Wiley and Sons, Inc., 1960.

[5] Moon, J.W. "Topics on Tournaments". New York, Holt, Rinehart and Winston, 1968.

[6] Narayana, T.V. and Zidek, J. "Contributions to the Theory of Tournaments", Part I, The Combinatorics of Knock-Out Tournaments, to appear in 'Cahiers du Bureau Universitaire de Recherche Opérationnelle Université de Paris', 1969.

[7] Narayana, T.V. and Zidek, J. "Contributions to the Theory of Tournaments", Part II, Statistical Inference in Random Tournaments, to appear in 'Revue Roumaine de Mathématiques pures et appliquées' 14, 1969.

[8] "A Million Random Digits with 100,000 Normal Deviates". The Rand Corporation. Glencoe, Illinois, The Free Press, 1955.